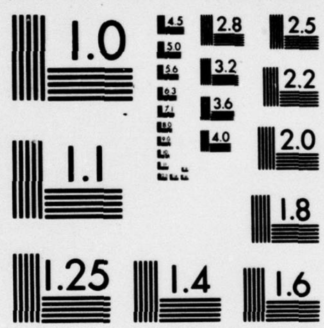
AD-A068 995 BROWN UNIV PROVIDENCE R I DIV OF APPLIED MATHEMATICS F/G 12/1
A ROBUST DISCRETE STATE APPROXIMATION TO THE OPTIMAL NONLINEAR --ETC(U)
MAR 79 H J KUSHNER N00014-76-0279
UNCLASSIFIED AFOSR-TR-79-0520 NL

1 OF 1
AD
A068 995

END
DATE
FILMED
7 — 79
DDC



MICROCOPY RESOLUTION TEST CHART

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD A068995

DDC FILE COPY

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER (18) AFOSR (19) TR- 79-0520	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) A ROBUST DISCRETE STATE APPROXIMATION TO THE OPTIMAL NONLINEAR FILTER FOR A DIFFUSION.		5. TYPE OF REPORT & PERIOD COVERED (9) Interim rept.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Harold J./Kushner (10)		8. CONTRACT OR GRANT NUMBER(s) (15) N00014-76-0279, AFOSR-76-3063
9. PERFORMING ORGANIZATION NAME AND ADDRESS Brown University Division of Applied Mathematics Providence, Rhode Island 02912		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F (16) 2304/A1 (17) A1
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM Bolling AFB, Washington, DC 20332		12. REPORT DATE (11) March 1979
		13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

LEVEL II

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

(12) 18 p.

DDC RECEIVED MAY 24 1979 C

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

ACCESSION for NTIS White Section DDC Buff Section UNANNOUNCED JUSTIFICATION BY DISTRIBUTION/AVAILABILITY CODES Dist. AVAIL. and/or SPECIAL A

sigma

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A robust computable approximation to the nonlinear filtering problem for a diffusion model is treated, where the system and data models are given by $dx = f(x)dt + \sigma(x)dz$, $dy = g(x)dt + dw$. The approximation (with approximation parameter h) is robust in the sense that it is locally Lipschitz continuous in the data $y(\cdot)$ (sup norm) uniformly in h and, as $h \to 0$, it converges to the optimal filter for the diffusion. approaches

DD FORM 1 JAN 73 1473 (Edit)

065 300 UNCLASSIFIED

AFOSR-TR- 79-0520

2

A ROBUST DISCRETE STATE APPROXIMATION TO THE OPTIMAL NONLINEAR FILTER FOR A DIFFUSION+

Harold J. Kushner
Division of Applied Mathematics and Engineering
Brown University
Providence, Rhode Island 02912

March, 1979

Abstract

A robust computable approximation to the nonlinear filtering problem for a diffusion model is treated, where the system and data models are given by $dx = f(x)dt + \sigma(x)dz$, $dy = g(x)dt + dw$. The approximation (with approximation parameter h) is robust in the sense that it is locally Lipschitz continuous in the data $y(\cdot)$ (sup norm) uniformly in h and, as $h \to 0$, it converges to the optimal filter for the diffusion.

+This research was supported in part by the Air Force Office of Scientific Research under AF-AFOSR 76-3063, in part by the National Science Foundation under MCS-76-07247-A-03 and in part by the Office of Naval Research under N0014-76-C-0279-P0002.

1. Introduction+

Let $z(\cdot), w(\cdot)$ denote two independent Wiener processes, the second one with covariance $I \cdot t$ and define the processes $x(\cdot), y(\cdot)$ by the Itô equations

$$dx = f(x)dt + \sigma(x)dz, \quad x(0) \text{ arbitrary } \varepsilon R^r, \tag{1.1}$$

$$dy = g(x)dt + dw, \quad y(0) = 0,\ y(t) \varepsilon R^s,\ t \leq T. \tag{1.2}$$

where T is an arbitrary, but fixed positive number.

It is assumed for convenience that $f(\cdot), \sigma(\cdot), g(\cdot)$ are bounded and continuous and the solution on $C^r[0,T]$ (the space of R^r-valued continuous functions on $[0,T]$) of (1.1) is unique in the sense of distributions. It is also assumed that the function $g(\cdot)$ has bounded and continuous first and second derivatives and either that the range of $x(\cdot)$ is bounded or that the derivatives of $g(\cdot)$ are uniformly continuous. Concerning the first condition, see the remarks in Section 4. We use x_t and $x(t)$ interchangeably.

Let $F(\cdot)$ denote a bounded continuous real valued function on R^r, and $\mathcal{Y}_t$ the σ-algebra generated by y_s, $s \leq t$. We are concerned with robust approximations to the conditional expectation $E_t F(x_t) \equiv E[F(x_t) | \mathcal{Y}_t]$, in the sense that we want a "good" approximation which is a continuous function of the data $y(\cdot)$.

Let $\bar{x}(\cdot)$ denote a process which is independent of $y(\cdot)$, but which induces the same measure on $C^r[0,T]$ that $x(\cdot)$ does. Then it is well known [1], [2] that w.p.1,

---

+The author gratefully acknowledges several stimulating discussions with J.M.C. Clark, who contributed considerably to the author's understanding of the subject.

$$(1.3)\qquad E_tF(x_t) = \frac{E_{\mathcal{Y}_t}F(\bar{x}_t)\exp[\int_0^t g'(\bar{x}_u)dy_u - \frac{1}{2}\int_0^t|g(\bar{x}_u)|^2du]}{E_{\mathcal{Y}_t}\exp[\int_0^t g'(\bar{x}_u)dy_u - \frac{1}{2}\int_0^t|g(\bar{x}_u)|^2du]} .$$

$E_{\mathcal{Y}_t}$ denotes expectation over $\bar{x}(\cdot)$ given $y(\cdot)$.

Although stochastic differential equation representations [1], [2] for $E_tF(x_t)$ are known, the problem of effectively computing or approximating $E_tF(x_t)$ is still in bad shape. Of particular interst is a computational or approximation method that is robust in $y(\cdot)$, in the sense that it is suitably continuous in $y(\cdot)$, and actually approximates $E_tF(x_t)$ well. Here, we develop an interesting approach to this problem by combining the approximation ideas in [3], [4] with the robustness ideas of Clark in [5]. In particular, Clark showed that (1.4) is also a version of $E_tF(x_t)$ and that it is locally Lipschitz continuous in $y(\cdot)$ at each $y(\cdot) \equiv C^1[0,T] \equiv C[0,T]$.

$$(1.4)\qquad E_tF(x_t) = \frac{E_{\mathcal{Y}_t}F(\bar{x}_t)\exp[y_t'g(\bar{x}_t) - \int_0^t y_u'dg(\bar{x}_u) - \frac{1}{2}\int_0^t|g(\bar{x}_u)|^2du]}{E_{\mathcal{Y}_t}\exp[y_t'g(\bar{x}_t) - \int_0^t y_u'dg(\bar{x}_u) - \frac{1}{2}\int_0^t|g(\bar{x}_u)|^2du]} .$$

In a formal sense, (1.4) is obtained from (1.3) by doing an integration by parts on the stochastic integral in (1.3).

In [3], Kushner developed a computational approach to optimal control and filtering problems on diffusion models. The basic idea was to approximate the diffusion in a particular way by an interpolated discrete parameter Markov chain, and to show that the minimal cost for the controlled chain converged to that for the diffusion, as some approximation parameter went to zero. For the filtering problem the filter for the approximating chain (but using the actual observational data from (1.2)) similarly converged to the filter for the diffusion. The interpolation times were constant in the filtering problem in [3]. We can use the same approximating process here as was used in [3]. In fact the filter approximation of [3] can be proved to be robust in the sense of Theorem 1 also. But it is more convenient notationally to use the continuous parameter Markov chain approximation $x^h(\cdot)$ to $x(\cdot)$ which was developed by Kushner and DiMasi ([4], Section 8) (h is an approximation parameter, see below).

Next, let $x^h(\cdot)$ denote <u>any</u> finite state continuous parameter Markov chain, and let us consider the corresponding filtering problem. Let the observational data available at t still be denoted by y_u, $u \le t$, where

$$(1.5) \qquad dy = g(x_t^h)dt + dw.$$

Then it is well known ([6], [7]) that (1.3) holds w.p.1 with $\bar{x}^h(\cdot)$, $x^h(\cdot)$ replacing $\bar{x}(\cdot)$, $x(\cdot)$ resp., where $\bar{x}^h(\cdot)$ has the same distributions as $x^h(\cdot)$ has, but is independent of $y(\cdot)$; i.e.,

$$(1.6) \quad E_t F(x_t^h) = \frac{E_{\mathscr{Y}_t} F(\bar{x}_t^h)\exp[\int_0^t g'(\bar{x}_u^h)dy_u - \frac{1}{2}\int_0^t |g(\bar{x}_u^h)|^2 du]}{E_{\mathscr{Y}_t} \exp[\int_0^t g'(\bar{x}_u^h)dy_u - \frac{1}{2}\int_0^t |g(\bar{x}_u^h)|^2 du]}.$$

Due to the piecewise constant nature of $\bar{x}^h(\cdot)$, the stochastic integral can be readily integrated by parts to yield the chain version of (1.4):

$$(1.7)\quad E_t F(x_t^h) = \frac{E_{\mathfrak{Y}_t} F(\bar{x}_t^h)\exp[g'(\bar{x}_t^h)y_t - \int_0^t y_u' dg(\bar{x}_u^h) - \frac{1}{2}\int_0^t |g(\bar{x}_u^h)|^2 du]}{E_{\mathfrak{Y}_t} \exp[g'(\bar{x}_t^h)y_t - \int_0^t y_u' dg(\bar{x}_u^h) - \frac{1}{2}\int_0^t |g(\bar{x}_u^h)|^2 du]}\ .$$

The expression (1.7) is also ([5]) locally Lipschitz continuous in $y(\cdot)$ at each $y(\cdot) \in C^s[0,T]$. In Section 2, we define a particular chain for which (1.7) can be computed and and also converges to (1.4) for all $y(\cdot)$. In Section 3, we prove the uniform robustness result; namely that the continuity in $y(\cdot)$ of (1.7) is uniform in h. This uniformity is crucial for robustness. Some remarks on computation are made in Section 4. Let $F_t^h(y(\cdot))$ and $F_t(y(\cdot))$ denote the value of (1.7) and (1.4), resp., at $y(\cdot)$.

2. A Useful Approximating Chain $\{x^h(\cdot)\}$.

We next describe the particular approximating chain which is to be used in Section 3. Other approximating chains can certainly be used. The main criteria for the choice concern the robustness and the ease of use in computation for the filtering problem. Let R_h^r denote the finite difference grid on R^r with difference parameter h. Either R_h^r or a subset of it will be our state space. Actually, h can be vector valued, the finite difference interval depending on the direction, but we stick to the simpler case. The transition function will be stated for the case where $a(\cdot) = \sigma(\cdot)\sigma'(\cdot)/2$ is diagonal, for notational simplicity. The expressions for the general case are in [3, Chapter 6.2]. Define $Q_h(x) = 2\sum_i a_{ii}(x) + h\sum_i |f_i(x)|$, $\Delta t^h(x) = h^2/Q^h(x)$, let $\inf_x(|a(x)| + |f(x)|) > 0$ and e_i = unit vector in i^{th} coordinate direction. For $x \in R_h^r$ and $y = x \pm e_i h$, set $p^h(x,y) = [a_{ii}(x) + hf_i^{\pm}(x)]/Q_h(x)$ where $f^+ = \max[0,f]$, $f^- = \max[0,-f]$; for other (x,y)-pairs, set $p^h(x,y) = 0$. Let $\{\xi_n^h\}$ denote the chain with transition probabilities $\{p^h(x,y)\}$. Interpolate the chain into a continuous parameter Markov process, denoted by $x^h(\cdot)$, by defining the interjump intervals by

$$P\{\text{jump after } t + s \mid x_t^h = x\} = \exp - (s/\Delta t^h(x)).$$

The process has the following properties [4, Section 8].

$$E[\text{next state value} - x \mid x = \text{current state value}] = f(x)\Delta t^h(x),$$

$$\text{covar}[\text{next state value} - x \mid x = \text{current state value}] = 2a(x)\Delta t^h(x) + o(\Delta t^h(x))$$

$$E[x^h_{t+\Delta} - x | x^h_t = x] = f(x)\Delta + o(\Delta),$$

$$\mathrm{covar}[x^h_{t+\Delta} - x | x^h_t = x] = 2a(x)\Delta + o(\Delta).$$

More details are in [3], [4]. The sequence $\{x^h(\cdot)\}$ converges weakly in $D^r[0,T)$ to the $x(\cdot)$ of (1.1), where $D^r[0,T]$ is the space of R^r valued functions on $[0,T]$ which have left-hand limits and are right continuous. The space is endowed with the Skorokhod topology. The states of the process $\{x^h(\cdot)\},\{\xi^h_n\}$ only communicate with their nearest neighbors.

Next, the main robustness theorem will be proved.

3. The Robustness Theorem.

Theorem 1. $E_tF(x_t^h)$ given by (1.7) is continuous in the supremum norm at each $y(\cdot) \in C^s[0,T]$, uniformly in (small) $h > 0$. In fact, for each bounded set S in $C^s[0,T]$, there is a real $K(S)$ such that $|F_t^h(y(\cdot)) - F_t^h(\tilde{y}(\cdot))| \leq K(S)||y - \tilde{y}||$ for $y(\cdot)$ and $\tilde{y}(\cdot) \in S$. Also, (1.7) converges to (1.4) for each $y(\cdot) \in C^s[0,T]$.

Proof. Let S denote a bounded set in $C^s[0,T]$, with $y(\cdot)$ and $\tilde{y}(\cdot) \in S$. We need only show that there is a constant $K_1(S)$ depending only on S such that

$$
\begin{aligned}
E\Big|&\exp[y_t'g(x_t^h) - \int_0^t y_u'dg(x_u^h) - \frac{1}{2}\int_0^t |g(x_u^h)|^2du] \\
&- \exp[\tilde{y}_t'g(x_t^h) - \int_0^t \tilde{y}_u'dg(x_u^h) - \frac{1}{2}\int_0^t |g(x_u^h)|^2du]\Big| \leq K_1(S)||y-\tilde{y}||.
\end{aligned}
\tag{3.1}
$$

uniformly in h and $t \leq T$. We can and will drop the $-\frac{1}{2}\int_0^t |g(x_u^h)|^2du$ term, since $g(\cdot)$ is bounded. Then, using the inequality $|e^x - e^y| \leq |x-y|(e^x+e^y)$, we have the upper bound for (3.1)

$$
\begin{aligned}
E|(y_t-\tilde{y}_t)'g(x_t^h) - \int_0^t (y_u-\tilde{y}_u)'dg(x_u^h)||&\exp[y_t'g(x_t^h) - \int_0^t y_u'dg(x_u^h)] \\
&+ \exp[\tilde{y}_t'g(x_t^h) - \int_0^t \tilde{y}_u'dg(x_u^h)]|.
\end{aligned}
$$

We need only show (3.2) and (3.3). K_2 is an arbitrary constant.

$$(3.2)\quad E\left|\int_0^t y_u' dg(x_u^h)\right|^2 \le K_2 ||y||^2, \text{ uniformly in } t \le T \text{ and small } h.$$

$$(3.3)\quad E \exp \int_0^t q_u' dg(x_u^h) \text{ bounded uniformly in bounded sets of } q(\cdot) \in C^s[0,T] \text{ and in (small) } h \text{ and } t \le T.$$

First, (3.2) will be proved. Let $\mathcal{B}_t^h$ and $\mathcal{B}(T)$ denote the minimal σ-algebra over which $g(x_u^h)$, $u \le t$, is measurable and the Borel field over $[0,T]$, resp. It is convenient to use the decomposition

$$(3.4)\quad g(x_t^h) = M_t^h + \Gamma_t^h,$$

where $M^h(\cdot)$ is a martingale and $\Gamma^h(\cdot)$ is a predictable process; in particular, $\Gamma^h(t)$ is adapted to $\mathcal{B}_t^h$ and (as an (ω,t) function), $\Gamma^h(\cdot)$ is measurable on the sub σ-algebra of $\mathcal{B}_T^h \times \mathcal{B}(T)$ which is induced by the left continuous functions. The decomposition (3.4) is unique and $\Gamma^h(\cdot)$ has the representation $\Gamma_t^h = \int_0^t \bar{\gamma}_s^h ds$ where $\bar{\gamma}_s^h = \gamma^h(x_s^h)$ and $\gamma^h(\cdot)$ is given by

$$(3.5)\quad \gamma^h(x) = \sum_y [g(y) - g(x)] p^h(x,y)/\Delta t^h(x)$$

Note that $\Gamma^h(\cdot)$ is the unique predictable function which satisfies

$$E[\Gamma_{t+s}^h - \Gamma_t^h | \mathcal{B}_t^h] = E[g(x_{t+s}^h) - g(x_t^h) | \mathcal{B}_t^h],$$

from which $\Gamma^h(\cdot)$ and $\gamma^h(\cdot)$ can be determined. Using a truncated Taylor expansion on (3.5), and the properties of the chain which were given in Section 2, and the definition $g_{xx}(x)$ = Hessian of $g(\cdot)$ at x, we can write

$$(3.6)\quad \gamma^h(x) = \{[g_x'(x)f(x) + \text{trace } a(x)g_{xx}(x)]\Delta t^h(x) + o(\Delta t^h(x)\}/\Delta t^h(x)$$

$$= \mathcal{L}g(x) + \varepsilon_t^h,$$

where $\varepsilon_t^h \to 0$ as $h \to 0$ uniformly in $t \leq T$, and in x, and $\mathcal{L}$ is the differential generator of (1.1). In any case $|\gamma^h(x)|$ is uniformly bounded in (small) h and in x. Note that $M^h(\cdot)$ is bounded on $[0,T]$, since $\gamma^h(\cdot)$ and $g(x^h(\cdot))$ are.

Next, let us calculate the quadratic variation of $M^h(\cdot)$. This is the increasing (in the sense of positive definite matrices) matrix valued predictable function $\Lambda^h(\cdot)$ such that

$$M_t^h(M_t^h)' - \Lambda_t^h \equiv N_t^h$$

is a matrix valued martingale. We have $\Lambda_t^h = \int_0^t \bar{\lambda}_s^h ds$, where $\bar{\lambda}_s^h = \lambda^h(x_s^h)$ and $\lambda^h(\cdot)$ is given by

$$\lambda^h(x) = \sum_y [g(y) - g(x)[[g(y) - g(x)]'p^h(x,y)/\Delta t^h(x).$$

$\bar{\lambda}^h(\cdot)$ is obtained from the characterization of $\Lambda^h(\cdot)$ as being the unique predictable function such that, for all $\delta > 0$,

$$E[\Lambda^h_{t+\delta} - \Lambda^h_t | \mathcal{B}^h_t] = E[M^h_{t+\delta}(M^h_{t+\delta})' - M^h_t(M^h_t)' | \mathcal{B}^h_t]$$

$$= E[(M^h_{t+\delta} - M^h_t)(M^h_{t+\delta} - M^h_t)' | \mathcal{B}^h_t].$$

Note that $|\lambda^h(x)|$ is bounded uniformly in x and h.

Now, we are prepared to evaluate (3.2). Write

$$\int_0^t y'_u dg(x^h_u) = \int_0^t y'_u dM^h_u + \int_0^t y'_u \bar{\gamma}^h_u du.$$

Then (3.2) follows from the uniform boundedness of $\bar{\gamma}^h(\cdot)$ and $\lambda^h(\cdot)$ and the martingale inequality

$$E \max_{t \le T} \left| \int_0^t y'_u dM^h_u \right|^2 \le 4||y||^2 E \int_0^T d\Lambda^h_u = 4||y||^2 E \int_0^T \bar{\lambda}^h_u du$$

We now turn to (3.3). It is convenient to bound (3.3) under the assumption that $x^h(\cdot)$ is stopped after the N^{th} jump, where N is an arbitrary integer. The obtained bound will not depend on N. Fix t, set $t = n\delta$, where n is an integer and write (3.3) as

$$A^h = E \prod_{i=0}^{n-1} \exp \int_{[i\delta, i\delta+\delta)} q'_u dg(x^h_u).$$

Let $E^h_{i\delta}$ denote the expectation conditioned on $\mathcal{B}^h_{i\delta}$. Let $x = x^h_{i\delta}$. There is a function $o(\cdot)$ which can depend on N and h and on the modulus of continuity of $q(\cdot)$ on $[0,T]$, but is uniform in x and is such that

$$A_i^h \equiv E_{i\delta}^h \exp \int_{[i\delta,i\delta+\delta)} q_u' dg(x_u^h) \le (1 - \frac{\delta}{\Delta t^h(x)} + o(\delta))$$

(3.7)

$$+ E_{i\delta}^h \exp \bar{q}_{i\delta}'[g(x^+) - g(x)] \cdot \frac{\delta}{\Delta t^h(x)} + o(\delta),$$

where x^+ is the successor state to x, given one jump in $[i\delta,i\delta+\delta)$ and $\bar{q}_{i\delta}$ is the value of q_s at the jump time in $[i\delta,i\delta+\delta)$ if any. Expanding (3.7) yields

$$\begin{aligned} A_i^h \le (1-\delta/\Delta t^h(x)) &+ (\delta/\Delta t^h(x) + o(\delta))E_{i\delta}^h[1+\bar{q}_{i\delta}'(g(x^+) - g(x)) \\ &+ (\bar{q}_{i\delta}'(g(x^+) - g(x)))^2/2 \\ &+ o_1[(\bar{q}_{i\delta}'(g(x^+) - g(x)))^2] + o(\delta) \end{aligned}$$

where $o_1(y)/y \to 0$ as $y \to 0$. Thus for some K_1, independent of N,h and x, and an $o(\cdot)$ with the properties of the above $o(\cdot)$ function

$$A_i^h \le 1 + K_1\delta + o(\delta).$$

Substituting this into (3.7) and letting $\delta \to 0$ yields

$$A^h \le \exp K_1 T, \text{ independently of } h \text{ and } N.$$

The last assertion of the theorem can be proved in the same way that a similar assertion was proved in [3, Chapter 7.5] for an interpolation of a chain similar to $\{\xi_n^h\}$. Q.E.D.

<u>Remarks</u>. 1. Since (1.7) also converges to (1.3) for almost all $y(\cdot)$ (Wiener measure) the theorem provides another proof of Clarks representation (1.4) and its Lipschitz continuity.

2. The uniformity (equi Lipschitz continuity) in the theorem is crucial to the value of the result, for otherwise the "robustness" could well become less and less as $h \to 0$.

3. From an applications point of view, robustness is important since the conditional moments should be smooth functions of the data. Otherwise unaccounted for errors in measurement or errors in the numerical calculation might render the result meaningless. Furthermore, in applications $w(\cdot)$ is not usually a Wiener process, although it is convenient to use a filter designed under the assumption that it <u>is</u> a Wiener process. Then the robustness idea is that if $w(\cdot)$ is close to a Wiener process (in some pathwise sense), then the estimates would also be close to the estimates which would be obtained if $w(\cdot)$ were actually a Wiener process. We might lose information by not building a filter which considers the actual statistical structure of (non Wiener) $w(\cdot)$, but that filter would be much more complicated than the one which is optimal under the Wiener assumption.

4. Remarks on Computation

First, let $\Delta > 0$ and let us define a process to be denoted by $\xi^{h,\Delta}(\cdot)$ and which is essentially $x^h(\cdot)$, but altered so that the jumps occur only at times $i\Delta$, $i = 1,\ldots$. Let $\Delta \leq \Delta t^h(x)$, all x. In particular, set $\xi^{h,\Delta}(t) = \xi_i^{h,\Delta}$ on $[i\Delta, i\Delta+\Delta)$, where $\{\xi_i^{h,\Delta}\}$ is a Markov chain on the state space R_h^r with transition probabilities $p_{(x,y)}^{h,\Delta} = p_{(x,y)}^h(\Delta/\Delta t^h(x))$ for $y = x \pm e_i h$ and $p_{(x,x)}^{h,\Delta} = 1 - \sum_{y=x\pm e_i h} p_{(x,y)}^{h,\Delta}$. The process $\xi^{h,\Delta}(\cdot)$ converges weakly to $x(\cdot)$ as $h, \Delta \to 0$.

Let $\bar{\xi}^{h,\Delta}(\cdot)$ be independent of $y(\cdot)$ but have the same path distributions as $\xi^{h,\Delta}(\cdot)$ does. Then with $\bar{\xi}^{h,\Delta}(\cdot)$ replacing $\bar{x}^h(\cdot)$ in (1.6), it is shown in [3, Chapter 7.5] that (1.6) converges to (1.3) for almost all $y(\cdot)$ (Wiener measures),

The method used in the proof of the theorem can also be used with the approximation $\xi^{h,\Delta}(\cdot)$ replacing the approximation $x^h(\cdot)$. Then both $\Gamma^h(\cdot)$ and $\Lambda^h(\cdot)$ will still be predictable, but will be piecewise constant in the $[i\Delta, i\Delta+\Delta)$ intervals, and the theorem will continue to hold. Furthermore the difference between (1.6) evaluated with $x^h(\cdot)$, and with $\xi^{h,\Delta}(\cdot)$, goes to zero as $\Delta, h \to 0$, uniformly in bounded $y(\cdot)$ sets.

In order to have a computationally feasible method for getting values of (1.6), the state space must be finite. Let G be a closed hyper-rectangle in R^r and set $G_h = R_h^r \cap G$. Let $\tau = \inf\{t: x_t \in \partial G\}$, $\tau' = \inf\{t: x_t \notin G\}$, and let μ denote the measure of x_0. Suppose that ∂G, the boundary of G, is regular in the sense that $P_\mu\{\tau \cap T = \tau' \cap T\} = 1$. Let $\tilde{x}^h(\cdot), \tilde{x}(\cdot), \tilde{\xi}^{h,\Delta}(\cdot)$ denote the processes stopped on first exit from G. Then the theorem remains valid for these processes replacing $x^h(\cdot), x(\cdot), \xi^{h,\Delta}(\cdot)$ (then $\bar{x}(\cdot)$ and $\bar{\xi}^{h,\Delta}(\cdot)$ are also replaced by the stopped processes). With the use of the stopped process, the state space G_h is finite and the method of [3, Chapter 7.5] can be used with the $\xi^{h,\Delta}(\cdot)$ approximation. Alternatively, we can use a method of Clark. In [5], Clark gave a set of ordinary differential equations (not Itô equations) for realizing (1.7), and the solution of this set, when considered as a function of $y(\cdot)$, has the same robustness property as (1.7) has.

REFERENCES

[1] H.J. Kushner, "Dynamical equations for non-linear filtering", Journal of Differential Equations, 1970, 179-190.

[2] M. Fujisaki, G. Kallianpur, H. Kunita, "Stochastic differential equations for the non-linear filtering problem", Osaka, J. Math., 1972, 19-40.

[3] H.J. Kushner, <u>Probability Methods for Approximations in Stochastic Control and for Elliptic Equations</u>, Academic Press, 1977, New York.

[4] H.J. Kushner, G. DiMasi, "Approximations for functionals and optimal control problems on jump diffusion processes", J. Math. Anal. and Applic. 1978, <u>63</u>, 772-800.

[5] J.M.C. Clark, "The design of robust approximations to the stochastic differential equations of non-linear filtering", Proc. Teddington Conf. on Communication Theory and Stochastic Processes, in Comm. Systems and Random Process Theory, ed. J.K. Skwirznski. Sijthoff, Noordhoff, Alphen an der Rijn, 1978.

[6] W.M. Wonham, "Some applications of stochastic differential equations to optimal non-linear filtering", SIAM J. on Control, 1965, <u>2</u>, 347-369.

[7] R.S. Liptser, A.N. Shiryayev, <u>Statistics of Random Processes</u>, 1977, Springer, Berlin (translation of the 1974 Russian original.